IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received August 29, 2019, accepted September 11, 2019, date of publication September 17, 2019, date of current version October 1, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2941899

Lumen Contour Segmentation in IVOCT Based on N-type CNN

JUNJIE TANG[1], YISHA LAN[1], SIRUI CHEN[1], YONGSHUO ZHONG[1], CHENXI HUANG[2], YONGHONG PENG[3], (Member, IEEE), QINYUAN LIU[1,7], YONGQIANG CHENG[4], FEI CHEN[5], AND WENLIANG CHE[6]

[1]Department of Computer Science and Technology, Tongji University, Shanghai 201804, China
[2]School of Informatics, Xiamen University, Xiamen 361005, China
[3]Faculty of Computer Science, University of Sunderland, Sunderland SR6 0DD, U.K.
[4]Department of Computer Science and Technology, University of Hull, Hull HU6 7RX, U.K.
[5]Department of Cardiology, Shanghai Tongji Hospital, Tongji University, Shanghai 200065, China
[6]Department of Cardiology, School of Medicine, Shanghai Tenth People's Hospital, Tongji University, Shanghai 200072, China
[7]Shanghai Institute of Intelligent Science and Technology, Tongji University, Shanghai, China

Corresponding authors: Qinyuan Liu (liuqy@tongji.edu.cn), Yongqiang Cheng (Y.Cheng@hull.ac.uk), Fei Chen (riverapt@126.com), and Wenliang Che (chewenliang@tongji.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61803282 and Grant 81570436, in part by the Fundamental Research Funds for the Central Universities of China under Grant 22120190211 and Grant 22120180561, and in part by the "Chenguang Program" supported by Shanghai Education Development Foundation and Shanghai Municipal Education Commission under Grant 18CG18.

ABSTRACT Automatic segmentation of lumen contour plays an important role in medical imaging and diagnosis, which is the first step towards the evaluation of morphology of vessels under analysis and the identification of possible atherosclerotic lesions. Meanwhile, quantitative information can only be obtained with segmentation, contributing to the appearance of novel methods which can be successfully applied to intravascular optical coherence tomography (IVOCT) images. This paper proposed a new end-to-end neural network (N-Net) for the automatic lumen segmentation, using multi-scale features based deep neural network, for IVOCT images. The architecture of the N-Net contains a multi-scale input layer, a N-type convolution network layer and a cross-entropy loss function. The multi-scale input layer in the proposed N-Net is designed to avoid the loss of information caused by pooling in traditional U-Net and also enriches the detailed information in each layer. The N-type convolutional network is proposed as the framework in the whole deep architecture. Finally, the loss function guarantees the degree of fidelity between the output of proposed method and the manually labeled output. In order to enlarge the training set, data augmentation is also introduced. We evaluated our method against loss, accuracy, recall, dice similarity coefficient, jaccard similarity coefficient and specificity. The experimental results presented in this paper demonstrate the superior performance of the proposed N-Net architecture, comparing to some existing networks, for enhancing the precision of automatic lumen segmentation and increasing the detailed information of edges of the vascular lumen.

INDEX TERMS IVOCT image, convolution neural network, cross entropy loss function, automatic segmentation.

I. INTRODUCTION

Coronary heart disease has become a widespread health concern causing fatality worldwide, also known as a myocardial ischemic heart disease caused by coronary artery stenosis or obstruction. As the degree of coronary stenosis continues to increase, blood vessels are restricted in blood transfusion functions, which may cause other symptoms such as angina. Therefore, segmentation [1] and quantitative analysis of the vascular lumen contour have significant contributions to the diagnosis of coronary heart disease. According to the degree of stenosis of the vascular lumen, coronary heart disease can be graded. It is also possible to optimize the size of the balloon and the size of the stent during the interventional procedure based on the segmented vessel contour information. As an intravascular imaging technique, intravascular

The associate editor coordinating the review of this manuscript and approving it for publication was Wenbing Zhao.

optical coherence tomography (IVOCT) can well show the shape and size of the lumen [2]. At present, the analysis of the contour of the vascular cavity mainly relies on manual segmentation by experts. However, each IVOCT usually pulls back 200 to 300 pictures, and it takes an average of 5 hours to complete the division task. Therefore, it has become crucial to develop an automatic segmentation algorithm to assist IVOCT vessel contours analysis.

Several fully automated and semi-automated IVOCT vessel contour segmentation algorithms have been proposed in the literature, e.g. Typical A-line based threshold segmentation, graph cut algorithm, and Markov random field model-based segmentation algorithm. Su *et al.* [3] used the artificial neural network (ANN) method as the feature learning algorithm for the detection of the lumen and MA borders in intravascular ultrasound (IVUS) images. Sun*et al.* [4] proposed a new graph-based approach for segmentation of luminal and external elastic lamina (EEL) surface of coronary vessels in gated 20 MHz IVUS image sequences (volumes). Essa *et al.* [5] proposed a combined cross-sectional segmentation approach with longitudinal tracking in order to tackle various forms of imaging artifacts and to achieve consistent segmentation. Malcolm *et al.* [6] presented how highly variable nonlinear shape priors learned from training sets can be added to existing iterative graph cut methods for accurate and efficient segmentation of such objects. However, these algorithms often need to pre-set relevant parameters according to experience. When the contour of the vascular lumen is irregular or there is a related stent implantation and vascular lumen thrombosis, the segmentation result is often poor.

In view of the problems encountered in the above segmentation methods, we propose an automatic segmentation algorithm for vascular cavity contours using deep neural networks based on multi-scale input fusion. Inspired by the capital letter N shape of the proposed convolutional network structure, we refer to the proposed network framework as an N-type network, namely the N-Net. The proposed N-Net consists of three parts: 1) multi-scale input layer, 2) N-type convolution network layer, 3) cross-entropy loss function. To the best of our knowledge, this is the first attempt in literature to use the deep learning technique to tackle the IVOCT vessel contour segmentation task. A series of comparison experiments with traditional U-Net has demonstrated the accuracy and stability of our method.

The remaining part of the paper is organized as follows. The related work is introduced in Section II. The proposed framework and N-Net is detailed in Section III. The experimental results are analyzed in Section IV and we conclude the paper in Section V.

II. RELATED WORK

A. U-NET

The U-Net network is one of the most famous Convolutional Neural Network (CNN) architectures for semantic segmentation. It is commonly used in medical convolutional neural networks for image segmentation and can accept input of any size. The U-Net architecture has achieved outstanding performance on various diverse biomedical segmentation applications with significant success. The U-Net architecture has potential to be applied to tasks beyond medical imaging applications, to wider computer vision tasks [8], [9]. The network consists of two main parts: a convolutional coding and decoding unit which performs the basic convolution operation and then activates ReLU in both parts of the network [10]. The U-Net model offers several advantages for segmentation tasks: Firstly, the model allows for the use of global locations and contexts simultaneously. Secondly, it requires very few training samples and provides good performance for segmentation tasks [8]. Thirdly, the end-to-end pipeline processes the entire image during the forward pass and directly produces the segmentation map. This ensures that U-Net retains the full context of the input image, which is a major advantage over patch-based segmentation [9]. The traditional U-Net network [7] simply splices the feature map of the encoder in the downsampling stage to the feature map of the upsampling stage to form a ladder network structure to obtain richer features. Since the detection of the contour of the vascular lumen requires a large amount of detailed information, as the network deepens, the detailed information is gradually lost. In addition, in the traditional U-Net network, feature fusion directly connects the feature maps [11]. This will result in no interaction between the feature maps and cannot be effectively merged.

B. CONVOLUTIONAL NEURAL NETWORK

The CNN [12] has become one of the most powerful tools in the field of machine learning and computer vision. These models have the ability to learn layered features from raw image data, so no handcrafted features are needed. In addition to the success of natural image recognition tasks such as image classification [13], CNN also showed good performance in various medical image calculation problems. Semantic segmentation is considered a dense pixel classification task that assigns labels to each pixel. Classification and positioning are often considered to be two major challenges in semantic segmentation. Models that are insensitive to location information can improve classification performance but reduce segmentation performance.

C. SEMANTIC SEGMENTATION

Semantic Segmentation [1] is a hot topic in the study of segmentation problems. Specifically, the goal of semantic segmentation is to assign all the pixels in the image to their corresponding labels (different from Object Detection and Localization). A general semantic partitioning architecture can be analogue to an encoder-decoder network. The encoder is usually a pre-trained classification network like VGG [19], AlexNet [13] and DenseNet [20], and then a decoder network. The difference between these architectures lies primarily in the decoder network. The task of the decoder is to semantically project discriminable features (lower resolution) learned

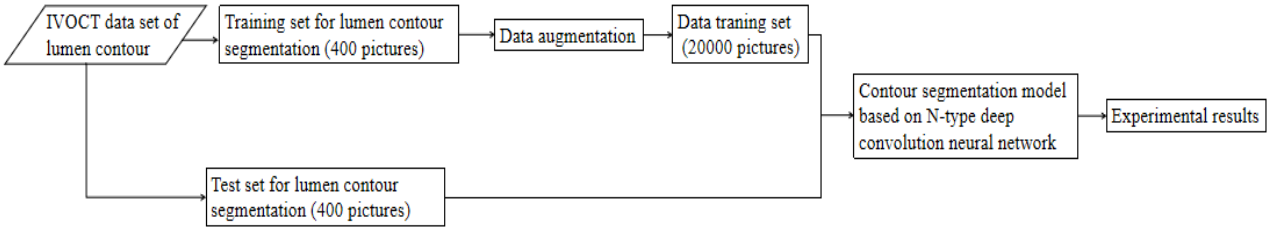


FIGURE 1. The overall flowchart of lumen contour segmentation based on N-Net.

by the encoder into pixel space (higher resolution) for dense classification.

III. METHOD

The overall flowchart of lumen contour segmentation is shown in Fig. 1. Firstly, 800 images labeled by experts are selected as experimental data. And among them, 400 images are selected as training set, while the remaining 400 images are used as test set. Secondly, considering the small amount of data, a variety of data augmentation methods are used to enhance the data training set of the network. And then 2000 images obtained from data augmentation are input to our proposed N-Net (N-type deep convolution neural network). Finally, a series of experiments are conducted to demonstrate the accuracy of the proposed N-Net.

Moreover, the architecture of proposed N-Net used in IVOCT lumen contour segmentation is shown in Fig. 2. As shown in Fig. 2, the U-Net network is contained as the main body of the proposed N-Net. The N-Net consists of three parts: the first part is multi-scale input layer, which is used to construct image pyramid input to achieve multi-level receptive field fusion; the second part is N-type convolution neural network, which can extract multi-level image features; the third part is the cross-entropy loss function, which is used to compare the difference between the experimental and manual results.

A. MULTI-SCALE INPUT LAYER

The traditional U-Net simply splices the feature graph of the coder from the down-sampling stage into the feature graph in the up-sampling stage to form a trapezoidal network structure to obtain richer features. With the deepening of network, a lot of detailed information required in contour detection is gradually lost. Hence the reusing of the original image as input can not only compensate for the information loss, but also fuse high and low resolution information to increase the generalization ability of the model. Thus, based on the fundamentals of the traditional U-Net, three feature graphs are used to fuse the N-Net in up-sampling, which are the feature graphs in down-sampling blocks, the feature graphs in up-sampling blocks and the original lumen images with various resolution sizes. The multi-scale feature graph fusion input in N-Net has two characteristics: 1) The details of feature information of each layer are increased through the input of multi-scale feature graph fusion. In the training process, the details are gradually lost with the deepening of layers of deep convolution network and the process of pooling operation. In order to retain as many detailed features as possible, the multi-scale feature graph fusion input can both enrich the detailed features of network and improve the accuracy and generalization ability of the mode. 2) In the traditional U-Net, feature fusion is a direct connection operation between feature graphs, which will lead to no information interaction between feature graphs and ineffective fusion. On this basis, if the new lumen images with different resolution are fused directly with the feature graph of N-Net fusion, the network parameters will be too much. For this purpose, the fusion is divided into two steps, and the dimension of fused feature graph is reduced to 1 by 1 convolution, contributing to reduction of the parameters and increment of the information interaction between features.

B. N-TYPE CONVOLUTION NEURAL NETWORK

N-type convolution neural network is the main body of the whole deep learning framework. The N-Net consists of feature extraction down-sampling network (left half of Fig. 2) and the up-sampling network (right half of Fig. 2). The feature extraction network follows the architecture of VGGNet-16 [1] convolution network, where 3 by 3 convolution kernel and 2 by 2 maximum pooling are used. After each layer of convolution operation, the Rectified Linear Unit (Relu) activation function is used to perform non-linear operation on extracted features. There are five down-sampling operations in feature extraction network. Up-sampling network is constructed by full convolution network U-Net and five up-sampling operations are included in total. Unlike the traditional U-Net, each up-sampling operation uses a feature graph that combines three features for up-sampling. After each up-sampling, the length and width of feature graph are doubled. Then the final result is the lumen contour, which can be considered as a two-label classification problem for segmentation of internal contour information. One is the point in the lumen contour, the other is the point outside the contour. Therefore, after last sampling, the size of output result should be the same as the original image. Meanwhile, in order to get the probability of each pixel belonging to the category, Sigmoid function is used in the last layer for activation, thus ensuring that the output probability value is between 0 and 1.

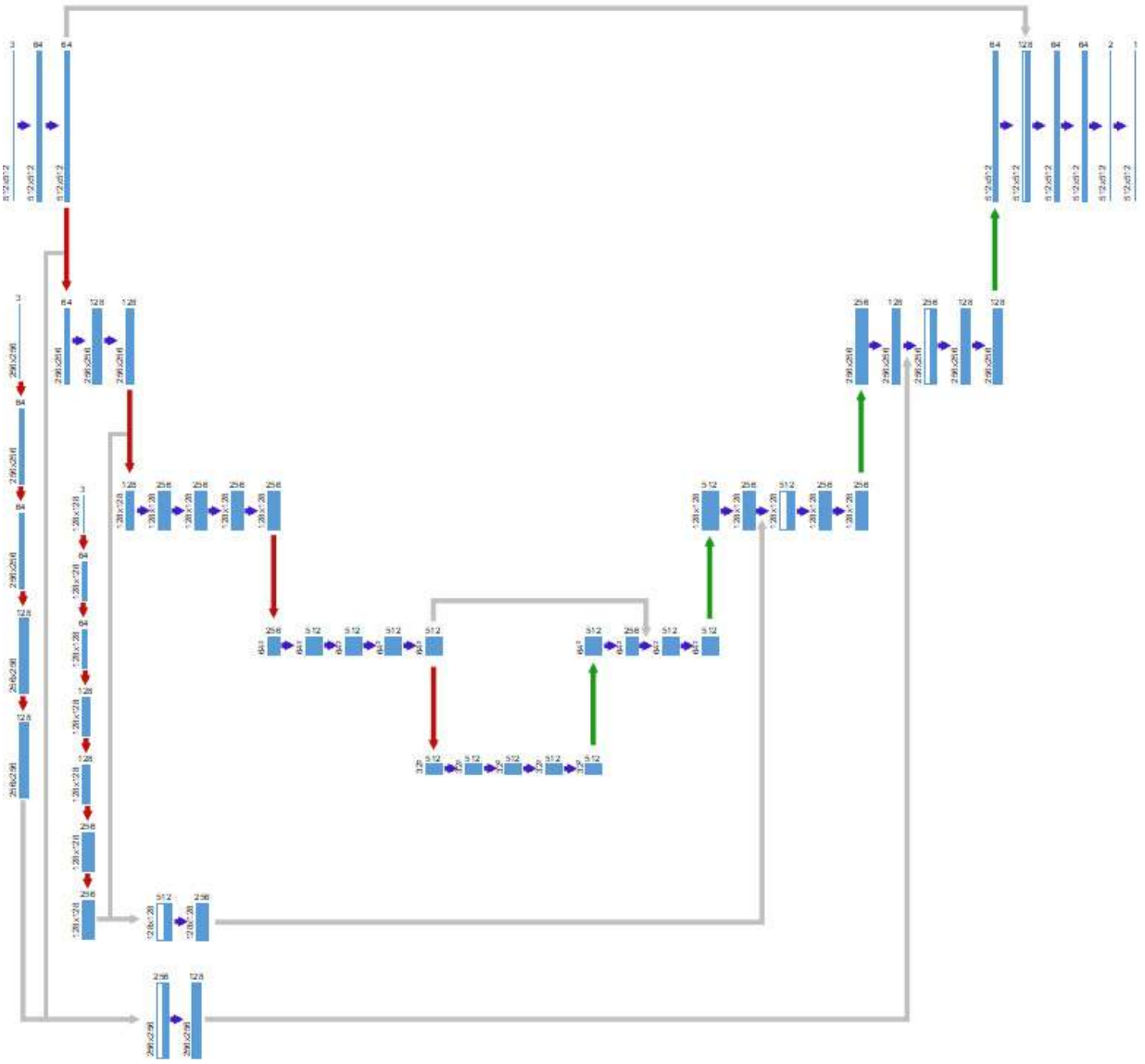


FIGURE 2. The architecture of N-Net, where multi-scale input layer of IVOCT, U-Net, output layer of lumen segmentation results and cross-entropy loss function are contained.

C. CROSS-ENTROPY LOSS FUNCTION

The cross-entropy loss function is chosen in the end-to-end contour segmentation. For any pixel x_{ij} in the input, the probability that the output pixel belongs to the contour lumen is represented as $p(w|x_{ij})$, and the cross-entropy function is defined as follows [1], [14]:

$$L(w) = -\frac{1}{N_i \times N_j} \sum_{i,j} [t_{ij} \ln p(w|x_{ij}) + (1 - t_{ij}) \ln(1 - p(w|x_{ij}))] \tag{1}$$

where $t_{ij} \in \{0, 1\}$ is the true category of a pixel x_{ij}, $t_{ij} = 1$ represents that the pixel is in the contour cavity while $t_{ij} = 0$ represents the pixel is outside the contour cavity.

D. STOCHASTIC GRADIENT DESCENT BASED ON MOMENTUM

In this part, the training set T with the size of N is defined as $T = \{(x_1, y_1), (x_2, y_2), \ldots, (x_N, y_N)\}$. And then the stochastic gradient descent based on momentum is mathematically given as following [15], [16]:

$$v_t = \gamma v_{t-1} - \eta \nabla_\theta J \sum_{i=1}^{m} J(y_i|x_i, \theta) \tag{2}$$

$$\theta' = \theta + v_t \tag{3}$$

where γ is the momentum term, η is the learning rate, ∇ is the gradient, t is the iteration time, v is the momentum, m is the number of random samples in each iteration and $J(\theta)$ represents the objective function. The gradient descent method based on momentum accumulates the

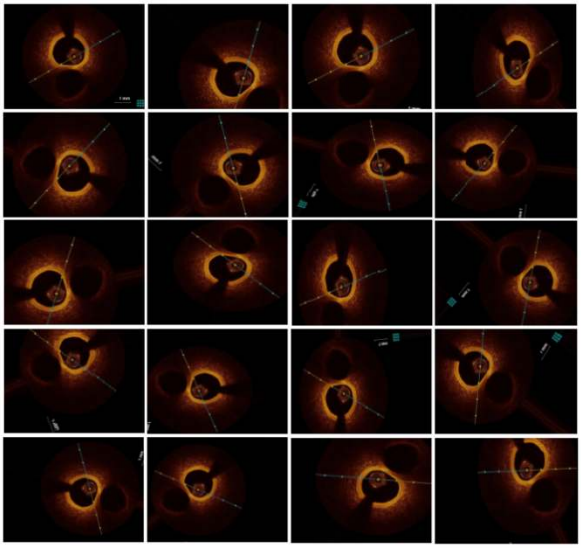


FIGURE 3. The contour results after data augmentation, where the image in first row and first column is the original data in training set, and the others are results of data augmentation.

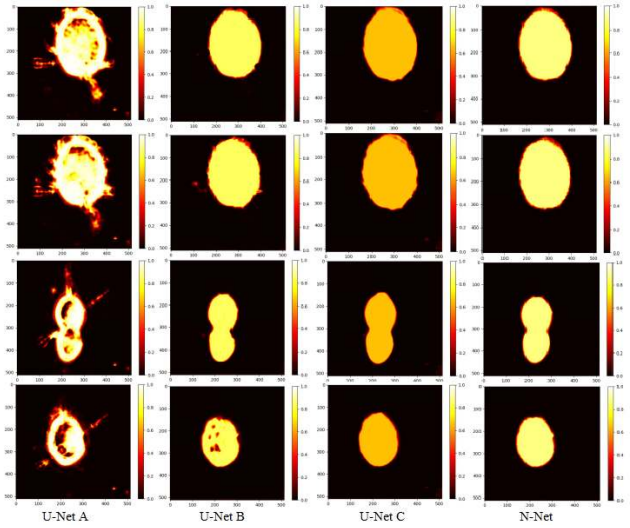


FIGURE 4. The probabilistic heat images of lumen pixels predicted by N-Net and different deep network, where the first two rows represent images with metal stent, the third row represents images with lateral branch, and the fourth row represents images with absorbable stent.

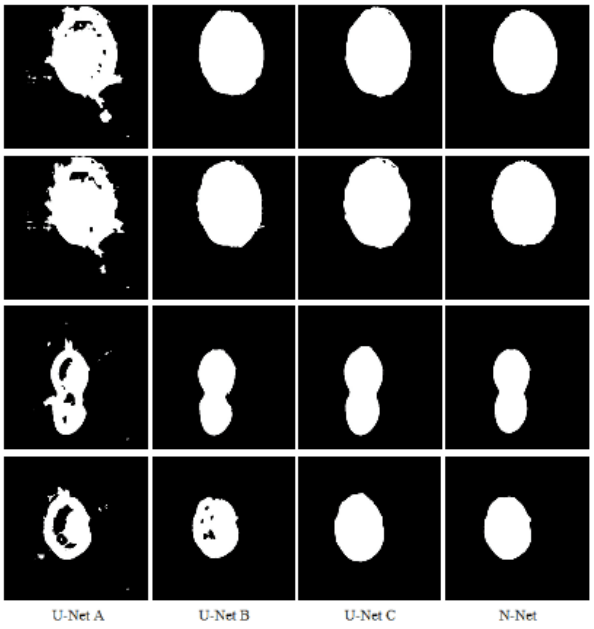


FIGURE 5. The binary lumen images on the basis of probability threshold 0.5, corresponding to Figure 4.

movement rate of the past gradient and can move faster in a direction when the gradient in the direction is uniform. The stochastic gradient descent based on momentum guarantees that the momentum term only updates the parameters of relevant samples so unnecessary parameter updates can be reduced. As a result, the approach achieves fast and stable convergence whilst effectively reduces the oscillation process.

IV. RESULT

A. EXPERIMENTAL ENVIRONMENT AND DATASET

The hardware configuration of our experimental platform is Intel(R) Xeon(R) 2.1 GHz CPU E5-2620, with 32GB memory. The NVIDIA TITAN X GPU is used to speed up convolutional neural network operation. The graphic memory is 12GB, its core frequency is 1417 MHz, and the single-precision floating-point operation ability is 11Tflops. In the experiment, Keras is used to build the N-Net and TensorFlow deep learning framework is adopted as the back-end framework of Keras. The procedure of blood vessel contour segmentation is written in Python and the experimental data are collected from Department of Cardiology, Oriental Hospital Affiliated to Tongji University. In the process of training, the training set contains 20,000 images. We set the size of each batch to 4, and train 4 images each time. A total of 10 epochs are trained and 5,000 iterations are performed. The stochastic gradient descent based on momentum is used to optimize the N-Net, where the momentum parameter is set to 0.8, the learning rate is set to 0.01, and the learning rate decay is 0.001. In data augmentation, several methods are introduced including rotation, scaling, distortion, mirror level turnover, translation, etc.. The results are shown in Fig. 3.

B. SUBJECTIVE EVALUATIONS OF EXPERIMENTAL RESULTS

The probabilistic heat images of lumen pixels predicted by N-Net and different deep networks is shown in Fig. 4. As the color becomes lighter, the probability that the point belongs to the contour increases. In the experiment, the first three blocks of VGGNet-16 is used as feature extraction network and designed as U-NetA. Similarly, the first four and the first five blocks of VGG are referred to as U-NetB and U-NetC respectively.

Fig. 5 shows the binary lumen image based on probability threshold. The white area is the predicted lumen area. We set

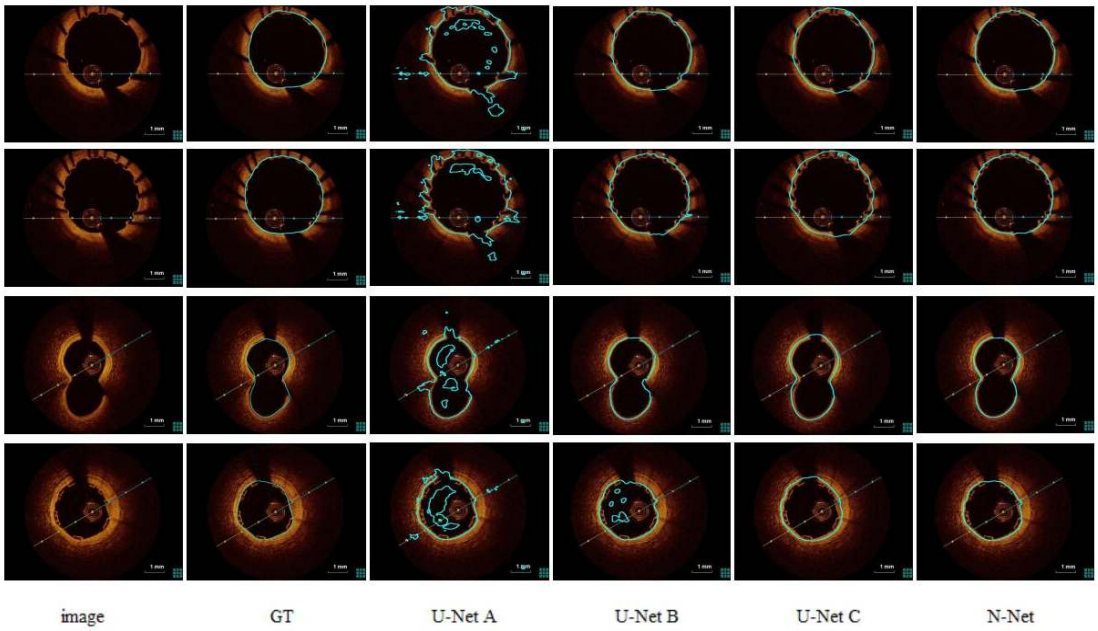


FIGURE 6. The segmentation results of lumen contour, where the first two rows represent images with metal stent, the third row represents images with lateral branch, and the fourth one represents images with absorbable stent.

the probability threshold to 0.5, when the probability value is greater than or equal to 0.5, we set it to 1, that is, the white area inside the vascular lumen; otherwise we set it to 0, that is, the black area outside the vascular lumen, so as to get the vascular lumen mask.

Moreover, the final results of lumen contour segmentation in IVOCT is shown in Fig. 6. In the segmentation results of U-Net A, it can be seen that the segmentation accuracy is seriously affected by the accumulation of thrombus in the vascular lumen. There are block-like missegmentation areas in the middle of the vascular lumen, and protruding burr-like areas in the edge of the vascular lumen, leading to the difficulty of showing the complete lumen contour. With the deepening of network layers, the missegmentation area of vascular thrombus gradually decreases, and the burr-like protuberance on the edge of the blood vessel become less obvious. Experiments show that with the deepening of network layers, the performance of U-Net segmentation is gradually improving. However, for the lumen images of metal stent and bio-absorbable stent, it can be seen that the lumen contour is affected by the shape of stent, and the shape of lumen is also irregular. As shown in Fig. 6, the proposed N-Net can effectively segment the vascular lumen images including metal stent, absorbable stent and lateral branch. The segmentation contour of N-Net is fairer and the missegmentation rate is low, contributing to effectively present the true shape of the vascular lumen contour.

C. OBJECTIVE EVALUATIONS OF EXPERIMENTAL RESULTS

A series of metrics are introduced to measure the accuracy of the proposed N-Net, considering the manual segmentation as a reference. The measures are given as follows [17], [18]:

TABLE 1. Segmentation results obtained by the proposed N-Net running five times randomly on verification set of IVOCT lumen contour.

Metric	Loss (%)	ACC (%)	Recall (%)	DSC (%)	JS (%)	Specificity (%)
Run 1	8.11	97.95	90.67	93.35	87.68	99.41
Run 2	8.11	97.95	90.41	93.22	87.43	99.41
Run 3	8.19	97.93	90.40	93.20	87.41	99.40
Run 4	7.87	98.01	90.74	93.41	87.77	99.41
Run 5	8.00	97.97	90.60	93.28	87.55	99.40
Average	8.06± 0.11	97.96 ±0.03	90.56 ±0.14	93.29 ±0.08	87.57 ±0.14	99.41 ± 0.005

1) Jaccard similarity coefficient (JS)

$$JS = \frac{|A \cap B|}{|A \cup B|}$$

2) Dice similarity coefficient (DSC)

$$DSC = 2 \cdot \frac{|A \cap B|}{|A| + |B|}$$

3) Accuracy (ACC)

$$ACC = \frac{TP + TN}{TP + FP + TN + FN}$$

TABLE 2. Comparative segmentation results between N-Net and different U-Net of vascular lumen contour.

Metric	Loss (%)	ACC (%)	Recall (%)	DSC (%)	JS (%)	Specificity (%)
U-Net A	10.75±0.14	96.30 ±0.05	85.11 ±0.26	87.26 ±0.12	77.53 ±0.19	98.34± 0.01
U-Net B	10.89±0.18	97.54 ±0.03	88.09 ±0.18	91.67 ±0.07	84.76 ±0.13	99.33± 0.01
U-Net C	14.54±0.20	97.43 ±0.03	**90.67 ±0.16**	91.63 ±0.06	84.76 ±0.11	98.81± 0.01
N-Net	**8.06 ±0.11**	**97.96 ±0.03**	90.56 ±0.14	**93.29 ±0.08**	**87.57 ±0.14**	**99.41± 0.005**

4) Recall

$$Recall = \frac{TP}{TP + FN}$$

5) Specificity

$$Specificity = \frac{TN}{TN + FP}$$

where A is the ground truth obtained by manual segmentation, and B is the predicted value generated by different neural networks. TP represents the predicted pixel classification results belonging to positive samples labelled manually. TN indicates that the predicted pixel classification results are negative agreeing with manual labels. FN represents the predicted pixel classification results are different from manual label, where the original positive samples are classified as negative samples. And FP indicates the predicted pixel classification results are different from manual labels, where original negative samples are classified as positive samples.

Fig. 7 shows the change of segmentation accuracy and loss of traditional U-Net for training set under different network layers. As shown in Fig. 7, with the increase of iterations, the cross-entropy loss function decreases gradually while the DSC, JS, ACC, Recall and Specificity are increasing. This indicates that with the deepening of U-Net, the effect of vascular lumen segmentation contour improves. Moreover, in the experiment, we use N-Net to segment testing data set data five times, and take the average of results of five times as the final segmentation result of N-Net. The experimental results are shown in Table 1. Meanwhile, the segmentation results between N-Net and different U-Net are shown in Table 2, demonstrating the better segmentation effects of N-Net.

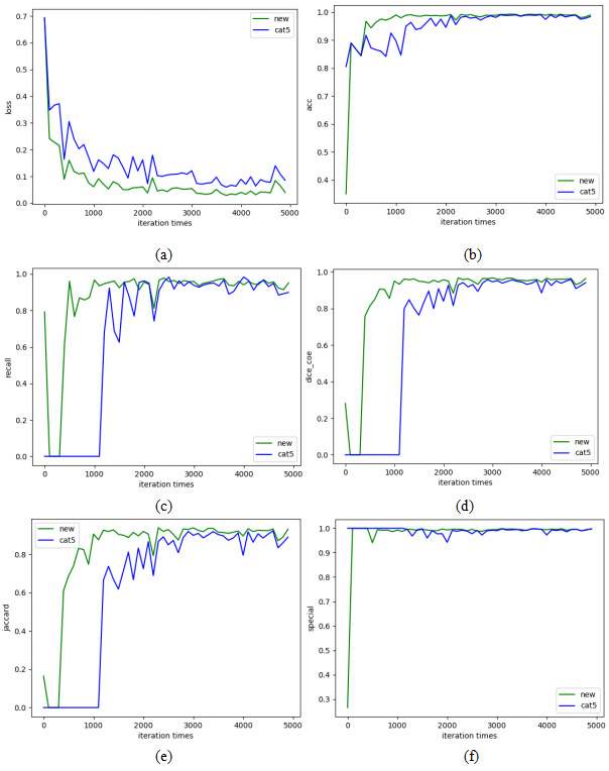


FIGURE 7. The segmentation result of between N-Net and U-Net C on IVOCT training with the increasing iterations, where (a) to (f) represent loss, accuracy, recall, DSC, JS and specificity respectively.

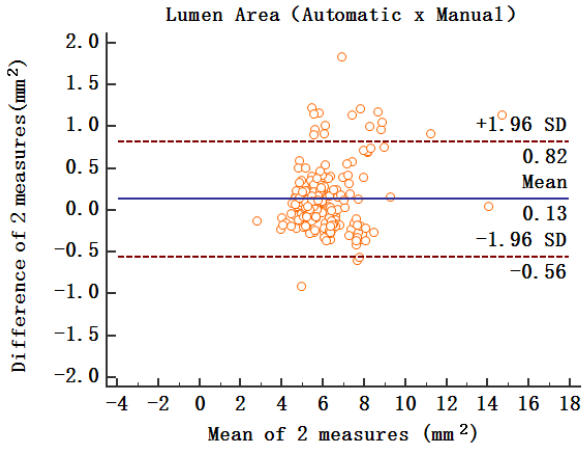


FIGURE 8. Bland-Altman diagram of N-Net segmentation and manual segmentation on IVOCT contour area analysis.

We randomly select 250 images and calculate the area size of each image contour segmented by the N-Net and manually, and then generate a Bland-Altman diagram to compare the consistency of the two segmentation methods. Fig. 8 shows the Bland-Altman diagram of the contour area obtained by automatic segmentation of N-Net and manual segmentation. As shown in Fig. 8, the scatter points are uniformly distributed in the standard deviation line, and the mean line is close to zero, demonstrating the similarity of the contour cavity area obtained by the above two methods. The correlation

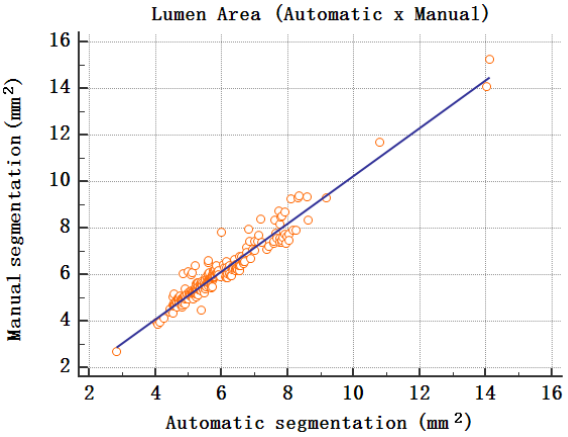


FIGURE 9. **Consistency analysis of N-Net segmentation and manual segmentation on IVOCT contour area.**

between the contour area obtained by automatic segmentation of N-Net and manual segmentation is shown in Fig. 9. The experimental results are r = 0.97, p<0.0001, which shows that the two methods have strong correlation.

V. CONCLUSION

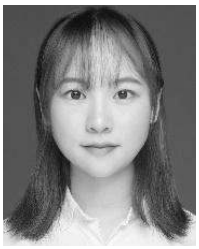
In this paper, a new deep learning framework named N-Net to segment IVOCT lumen contours is proposed. The N-Net model solves the challenges of segmenting images with stent implantation and thrombosis accumulation by fusing multi-scale feature maps of the vascular lumen. The multi-scale input layer enhances the detailed information of the vascular lumen. The N-Net is composed of five consecutive down-sampling operations and five consecutive up-sampling operations respectively. The selected cross entropy loss function ensures a high degree of matching between the actual output and the predicted output. The effect of N-Net on the selected IVOCT data set are validated and compared with the manually segmented results. The experimental results show that the proposed N-Net network can accurately segment the images containing vascular stents and thrombus compared with traditional U-Net, and the edges of the vascular lumen are smoother and contain more edge details. This automatic process will significantly reduce the workload of segmentation tasks required by doctors. In the future, new loss function will be explored for contour segmentation of vascular lumen to improve the accuracy of segmentation.

REFERENCES

[1] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2015, pp. 3431–3440.
[2] M. J. Suter, G. J. Tearney, W.-Y. Oh, and B. E. Bouma, "Progress in intracoronary optical coherence tomography," *IEEE J. Sel. Topics Quantum Electron.*, vol. 16, no. 4, pp. 706–714, Jul./Aug. 2010.
[3] S. Su, Z. Hu, Q. Lin, W. K. Hau, Z. Gao, and H. Zhang, "An artificial neural network method for lumen and media-adventitia border detection in IVUS," *Comput. Med. Imag. Graph.*, vol. 57, pp. 29–39, Apr. 2017.
[4] S. Sun, M. Sonka, and R. R. Beichel, "Graph-based IVUS segmentation with efficient computer-aided refinement," *IEEE Trans. Med. Imag.*, vol. 32, no. 8, pp. 1536–1549, Aug. 2013.
[5] E. Essa and X. Xie, "Automatic segmentation of cross-sectional coronary arterial images," *Comput. Vis. Image Understand.*, vol. 165, pp. 97–110, Dec. 2017. doi: 10.1016/j.cviu.2017.11.004.
[6] J. Malcolm, Y. Rathi, and A. Tannenbaum, "Graph cut segmentation with nonlinear shape priors," in *Proc. ICIP*, Sep./Oct. 2007, pp. 365–368.
[7] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.* Cham, Switzerland: Springer, 2015, pp. 234–241.
[8] Z. Zhang, Q. Liu, and Y. Wang, "Road extraction by deep residual U-net," 2017, *arXiv:1711.10684*. [Online]. Available: https://arxiv.org/abs/1711.10684
[9] R. Li, "DeepUNet: A deep fully convolutional network for pixel-level sea-land segmentation," 2017, *arXiv:1709.00201*. [Online]. Available: https://arxiv.org/abs/1709.00201
[10] M. Z. Alom, M. Hasan, C. Yakopcic, T. M. Taha, and V. K. Asari, "Recurrent residual convolutional neural network based on U-net (R2U-net) for medical image segmentation," 2018, *arXiv:1802.06955*. [Online]. Available: https://arxiv.org/abs/1802.06955
[11] Y. Han and J. C. Ye, "Framing U-net via deep convolutional framelets: Application to sparse-view CT," *IEEE Trans. Med. Imag.*, vol. 37, no. 6, pp. 1418–1429, Jun. 2018.
[12] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proc. IEEE*, vol. 86, no. 11, pp. 2278–2324, Nov. 1998.
[13] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Proc. Adv. Neural Inf. Process. Syst. (NIPS)*, vol. 2012, pp. 1097–1105.
[14] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Proc. 18th Int. Conf. Med. Image Comput. Comput.-Assisted Intervent. (MICCAI)*, in Lecture Notes in Computer Science, vol. 9351, Munich, Germany. Cham, Switzerland: Springer, Oct. 2015, pp. 234–241. doi: 10.1007/978-3-319-24574-4_28.
[15] N. Qian, "On the momentum term in gradient descent learning algorithms," *Neural Netw.*, vol. 12, no. 1, pp. 145–151, Jan. 1999.
[16] R. S. Sutton, "Two problems with backpropagation and other steepest-descent learning procedures for networks," in *Proc. Cogn. Sci. Soc.*, 1986.
[17] C. Huang, X. Shan, Y. Lan, L. Liu, H. Cai, W. Che, Y. Hao, Y. Cheng, and Y. Peng, "A hybrid active contour segmentation method for myocardial D-SPECT images," *IEEE Access*, vol. 6, pp. 39334–39343, 2018.
[18] X. Wang, Y. Lan, F. Su, Z. Li, C. Huang, Y. Hao, L. Yue, Y. Cheng, and Y. Peng, "Automatic stent strut detection in intravascular ultrasound using feature extraction and classification technique," in *Proc. 14th Int. Conf. Natural Comput., Fuzzy Syst. Knowl. Discovery (ICNC-FSKD)*, Jul. 2018, pp. 1172–1176.
[19] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, *arXiv:1409.1556*. [Online]. Available: https://arxiv.org/abs/1409.1556
[20] G. Huang, Z. Liu, L. van der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," 2016, *arXiv:1608.06993*. [Online]. Available: https://arxiv.org/abs/1608.06993

JUNJIE TANG is currently pursuing the master's degree with Tongji University, Shanghai, China. His research interests include image processing, reconstruction, and 3-D visualization.

YISHA LAN is currently pursuing the B.Sc. degree with Tongji University, Shanghai, China. Her research interests include image processing, reconstruction, and 3-D visualization.



SIRUI CHEN is currently pursuing the B.Sc. degree with Tongji University, Shanghai, China. Her research interests include machine learning, and image processing and reconstruction.

YONGSHUO ZHONG is currently pursuing the B.Sc. degree with Tongji University, Shanghai, China. His research interest includes image processing and reconstruction.

CHENXI HUANG received the B.Sc. degree in computer science from Tongji University, Shanghai, China, in 2015, where he is currently pursuing the Ph.D. degree in computer science. He is also an Assistant Professor with the School of Informatics, Xiamen University. His research interests include image processing, image reconstruction, data fusion, 3-D visualization, and machine learning.

YONGHONG PENG (M'02) is currently a Professor of Data Science (Chair) and the Head of Data Science Research, University of Sunderland, U.K. His research areas include data science, machine learning, data mining, and artificial intelligence. Prof Peng is the Chair for the Big Data Task Force (BDTF), and a member of Data Mining and Big Data Analytics Technical Committee of IEEE Computational Intelligence Society (CIS). He is also a founding Member of the Technical Committee on Big Data (TCBD) of IEEE Communications and an advisory board member for the IEEE Special Interest Group (SIG) on Big Data for Cyber Security and Privacy. He is an Associate Editor for the IEEE TRANSACTION ON BIG DATA, and an Academic Editor of PeerJ and PeerJ Computer Science.

QINYUAN LIU received the B.Eng. degree in measurement and control technology and instrumentation from the Huazhong University of Science and Technology, Wuhan, China, in 2012, and the Ph.D. degree in control science and engineering from Tsinghua University, Beijing, China, in 2017. He was a Visiting Student from September 2014 to March 2015 and a Research Assistant from July 2015 to September 2015, with the Department of Electronic and Computer Engineering, The Hong Kong University of Science and Technology, Hong Kong. From January 2016 to January 2017, he was an International Researcher with the Department of Computer Science, Brunel University London, U.K. He is currently an Assistant Professor with the Department of Computer Science and Technology, Tongji University, Shanghai, China. His research interests include machine learning, networked control systems, and distributed filtering. He is an Active Reviewer for many international journals.

YONGQIANG CHENG is currently a Senior Lecturer with the Department of Computer Science and Technology, University of Hull, U.K. His research interests include digital healthcare technologies, embedded systems, control theory and applications, artificial intelligence, and data mining.

FEI CHEN received the Ph.D. and M.D. degrees from the Medical College, Zhejiang University, China. He is currently a Fellow Doctor with Tongji Hospital, Tongji University, Shanghai. His research interests include the pathophysiological proceeding of atherosclerosis, interventional treatment of CHD, image reconstruction of coronary stent, and 3-D visualization of coronary stent.

WENLIANG CHE received the Ph.D. and M.D. degrees from Tongji University, Shanghai, China. He is currently an Interventional Cardiologist with the Department of Cardiology, School of Medicine, Shanghai Tenth People's Hospital, Tongji University, Shanghai. His research interests include atherosclerosis imaging, intra-coronary imaging in managing CHD, and quantitative assessment of coronary microvascular function.